



## Using CO<sub>2</sub> spatial variability to quantify representation errors of satellite CO<sub>2</sub> retrievals

A. A. Alkhaled,<sup>1</sup> A. M. Michalak,<sup>1,2</sup> and S. R. Kawa<sup>3</sup>

Received 30 April 2008; revised 26 June 2008; accepted 9 July 2008; published 29 August 2008.

[1] Satellite measurements of column-averaged CO<sub>2</sub> dry-air mole fraction ( $X_{\text{CO}_2}$ ) will be used in inversion and data assimilation studies to improve the precision and resolution of current estimates of global fluxes of CO<sub>2</sub>. Representation errors due to the mismatch in spatial scale between satellite retrievals and atmospheric transport models contribute to the uncertainty associated with flux estimates. This study presents a statistical method for quantifying representation errors as a function of the underlying spatial variability of  $X_{\text{CO}_2}$  and the spatial distribution of retrieved soundings, without knowledge of the true  $X_{\text{CO}_2}$  distribution within model gridcells. Representation errors are quantified globally using regional  $X_{\text{CO}_2}$  spatial variability inferred using the PCTM/GEOS-4 model and a hypothetical atmospheric transport model with  $1^\circ \times 1^\circ$  resolution, 3 km<sup>2</sup> retrieval footprints, and two different sounding densities. **Citation:** Alkhaled, A. A., A. M. Michalak, and S. R. Kawa (2008), Using CO<sub>2</sub> spatial variability to quantify representation errors of satellite CO<sub>2</sub> retrievals, *Geophys. Res. Lett.*, 35, L16813, doi:10.1029/2008GL034528.

### 1. Introduction

[2] Satellite missions, such as the Orbiting Carbon Observatory (OCO) and the Greenhouse Gases Observing Satellite (GOSAT), will provide global data of column-averaged CO<sub>2</sub> dry-air mole fraction ( $X_{\text{CO}_2}$ ) at high spatial resolutions. These data will be used in inverse modeling studies to improve the precision and resolution of current estimates of global carbon budgets [Rayner and O'Brien, 2001; Houweling et al., 2004; Chevallier et al., 2007]. The amount of information that satellite retrievals contribute towards improving CO<sub>2</sub> flux estimates will depend on their error characteristics; therefore, an accurate evaluation of the error statistics of retrieved soundings is central to providing accurate estimates of CO<sub>2</sub> sources and sinks and their associated uncertainties [Chevallier et al., 2007; Engelen et al., 2002].

[3] In inverse modeling studies, observation errors (a.k.a. model-data mismatch) are a combination of: (1) measurement errors due to the satellite instrument, and any approximations or errors in the retrieval algorithm, (2) transport model errors due to modeling simplifications and the uncertainties of model parameters, (3) aggregation errors

caused by estimating CO<sub>2</sub> fluxes at temporal and spatial resolutions coarser than the transport model, and (4) representation errors due to the resolution mismatch between observations and model gridcells [Enting, 2002; Engelen et al., 2002; Michalak et al., 2005]. Representation errors are attributed to the inability of atmospheric transport models to resolve the spatial and temporal variations captured by CO<sub>2</sub> observations, due to the low spatial and temporal resolution of the models relative to that of measurements [Engelen et al., 2002; Gerbig et al., 2003]. In theory, the concentration value assigned to a model gridcell should be equal to the true  $X_{\text{CO}_2}$  mean over the area of the gridcell and during the model time-step. In reality, the true mean is not known and is instead estimated from the satellite retrievals. The representation error is therefore equal to the uncertainty associated with the inferred gridcell mean, given the satellite retrievals over the gridcell, and is a function of  $X_{\text{CO}_2}$  variability over the sampled region. For example, sparse retrievals over a gridcell located in a region with high  $X_{\text{CO}_2}$  variability will be less likely to capture the true mean for that gridcell, and will have higher representation error.

[4] A number of studies have provided an evaluation of the representation error of observations used in inverse modeling studies. Rödenbeck et al. [2003] approximated representation errors using the standard deviation of simulated CO<sub>2</sub> concentrations of gridcells surrounding a measurement location. Although this approximation provides a measure of simulated CO<sub>2</sub> variability at the gridcell resolution in the region of a measurement, it does not evaluate the representativeness of a measurement of the mean CO<sub>2</sub> concentration within a gridcell.

[5] van der Molen and Dolman [2007] studied the representation error of measurements of CO<sub>2</sub> based on model simulations. Their analysis showed that representation errors increase with CO<sub>2</sub> variability. The study quantified these errors empirically as the average standard deviation of simulated CO<sub>2</sub> fields within different radii of measurement locations. This approach, however, requires knowledge of the entire sampled distribution (e.g.  $X_{\text{CO}_2}$  over gridcell) and does not evaluate the representativeness of multiple measurements within a given gridcell.

[6] Gerbig et al. [2003] and Lin et al. [2004] evaluated the spatial covariance of partial CO<sub>2</sub> columns using aircraft measurements. The studies used the evaluated spatial covariance to statistically generate simulated fields with a similar spatial covariance at small spatial resolutions. The simulated fields were divided into subareas used to represent model gridcells. The representation error was then evaluated as the average standard deviation of the simulated values within each model gridcell. This evaluation reflects the variance of the potential retrievals within a model gridcell, but does not represent the uncertainty in estimating

<sup>1</sup>Department of Civil and Environmental Engineering, University of Michigan, Ann Arbor, Michigan, USA.

<sup>2</sup>Department of Atmospheric, Oceanic and Space Sciences, University of Michigan, Ann Arbor, Michigan, USA.

<sup>3</sup>NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.

the gridcell mean given multiple measurements within each gridcell.

[7] In the context of satellite data, a number of studies have evaluated the representation error as the within-gridcell  $X_{CO_2}$  variance (or sampling variance). *Corbin et al.* [2008] and *Miller et al.* [2007] evaluated representation errors empirically based on high resolution  $X_{CO_2}$  simulations. *Miller et al.* [2007] sampled model gridcells according to a North-South swath, and assumed that the representation error is equal to the difference between the true simulated gridcell mean and the sample mean. *Corbin et al.* [2008] extended this approach to include temporal variability and the effect of clouds, by excluding cloudy pixels from the sampled North-South swaths. Both studies calculated the swath means of all possible swath locations within model gridcells, and used the statistics of the resulting swath mean distribution. *Corbin et al.* [2008] subtracted the known simulated gridcell means from these distributions, and used the standard deviation of the residuals as an estimate of the representation error. The methods presented in these two studies cannot be reproduced for actual satellite sampling conditions, however, because the true gridcell means are unknown.

[8] This study introduces a statistical method for evaluating the representation errors associated with using satellite retrievals to represent the mean  $X_{CO_2}$  within atmospheric transport model gridcells. The proposed method is based on: (1) the spatial distribution of satellite retrievals within a model gridcell, and (2) knowledge of the degree of  $X_{CO_2}$  variability in the vicinity of the model gridcell. The proposed method uses a geostatistical evaluation of the  $X_{CO_2}$  variability to quantify the spatial covariance between any two satellite retrievals as a function of their separation distance. This spatial covariance function can be inferred from available in situ data,  $X_{CO_2}$  model simulations, or potentially from the satellite retrievals themselves. Together with known retrieval locations, the method evaluates representation errors in a way that: (1) reflects the amount of information provided by available retrievals about the true unknown gridcell mean, and (2) does not require knowledge of the true value of that mean. The method is demonstrated using the regional spatial covariance statistics derived by *Alkhaled et al.* [2008] using modeled  $X_{CO_2}$ , together with assumed spatial distributions of satellite retrievals within hypothetical model gridcells.

## 2. Data and Methods

### 2.1. Methods

[9] When  $X_{CO_2}$  measurements are used as observations within a model, the  $X_{CO_2}$  value assigned to a given gridcell is intended to represent the true mean of the  $X_{CO_2}$  distribution within that gridcell. In reality, however, individual OCO  $X_{CO_2}$  soundings will have a much smaller footprint relative to a typical atmospheric transport model gridcell, and these soundings will not sample the full area of gridcells. Therefore, statistically, the representation error is the uncertainty associated with inferring the mean  $X_{CO_2}$  for a given gridcell using retrieved soundings. The proposed method evaluates representation errors using block kriging [e.g., *Chilès and Delfiner*, 1999], a spatial estimation method that uses the spatial covariance information of

$X_{CO_2}$  over sampled regions together with information about the locations of retrieved soundings to quantify the uncertainty associated with the mean  $X_{CO_2}$  within each model gridcell (i.e. representation error  $\sigma_{RE}$ ).

[10] To construct the block kriging system, each model gridcell is divided into  $m$  pixels with areas equal to the satellite sounding footprint (e.g. 3 km<sup>2</sup> for OCO). The retrievals are assumed to be an  $n \times 1$  vector of noisy samples  $\mathbf{z}$  taken at locations  $\mathbf{x}$  of a random spatial process  $\mathbf{y}$  representing the  $X_{CO_2}$  distribution within the gridcell at the resolution of satellite soundings:

$$\mathbf{z}(\mathbf{x}) = \mathbf{y}(\mathbf{x}) + \boldsymbol{\epsilon} \quad (1)$$

[11] The retrieval measurement errors ( $\boldsymbol{\epsilon}$ ) have an  $n \times n$  covariance matrix  $\mathbf{R}$ , which can be diagonal if the errors are assumed to be uncorrelated, or can have off-diagonal elements to represent spatially-correlated retrieval errors. The  $X_{CO_2}$  distribution within the gridcell at the resolution of satellite soundings ( $\mathbf{y}$ ) is assumed to have a mean  $E[\mathbf{y}] = \mathbf{X}\boldsymbol{\beta}$ , where  $\mathbf{X}$  is a matrix of covariate values at the sampling locations,  $\boldsymbol{\beta}$  is a vector of coefficients, and  $E[.]$  is the expectation operator. For the current application, the spatial mean ( $E[\mathbf{y}]$ ) within each gridcell is assumed constant (although it can vary between gridcells); therefore,  $\mathbf{X}$  is an  $m \times 1$  vector of ones and  $\boldsymbol{\beta}$  is an unknown large-scale mean.  $\mathbf{y}$  is also described using an  $m \times m$  spatial covariance matrix  $\mathbf{Q} = E[(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T]$ . Each element of the covariance matrix ( $Q_{ij}$ ) is calculated based on the regional spatial covariance and the separation distances ( $h_{ij}$ ) between the gridcell pixels. For example, for an exponential covariance structure, the elements of  $\mathbf{Q}$  will have the form [e.g., *Chilès and Delfiner*, 1999]:

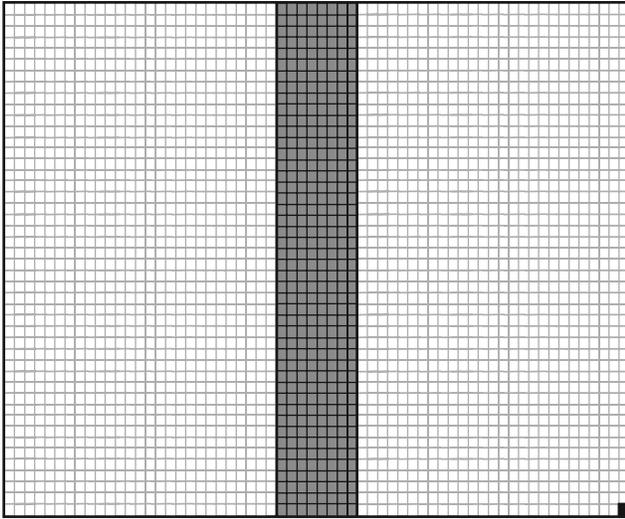
$$Q_{ij} = \sigma_{reg}^2 \exp\left(-\frac{h_{ij}}{L_{reg}}\right) \quad (2)$$

where  $\sigma_{reg}^2$  and  $L_{reg}$  represent the regional  $X_{CO_2}$  variance and range parameter, respectively, and where the distance beyond which the correlation between any two  $X_{CO_2}$  measurements approaches zero (i.e. the correlation length) is  $3L_{reg}$ .

[12] The uncertainty associated with the estimated  $X_{CO_2}$  distribution within a gridcell ( $\hat{\mathbf{y}}$ ) at the resolution of the satellite soundings can be quantified by solving the following kriging system:

$$\begin{bmatrix} \mathbf{S}\mathbf{Q}\mathbf{S}^T + \mathbf{R} & \mathbf{S}\mathbf{X} \\ (\mathbf{S}\mathbf{X})^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Lambda}^T \\ \mathbf{M} \end{bmatrix} = \begin{bmatrix} \mathbf{S}\mathbf{Q} \\ \mathbf{X}^T \end{bmatrix} \quad (3)$$

where  $\mathbf{S}$  is an  $n \times m$  indicator matrix of zeros and ones, with each row of  $\mathbf{S}$  corresponding to a single satellite retrieval, and a one indicating the location of the sampled pixel. Equation 3 is solved for an  $m \times n$  matrix of coefficients  $\boldsymbol{\Lambda}$  and a  $1 \times m$  vector of Lagrange multipliers  $\mathbf{M}$ . The  $\boldsymbol{\Lambda}$ 's represent the weighting that each of the  $n$  retrieved soundings receives in estimating the  $X_{CO_2}$  value at each of the  $m$  locations within the gridcell, and  $\mathbf{M}$  represents the additional uncertainty resulting from the fact that the mean of the spatial process  $\mathbf{y}$  is assumed unknown.



**Figure 1.** A  $1^\circ \times 1^\circ$  model gridcell at  $45^\circ$  latitude discretized into  $3\text{km}^2$  pixels representing the resolution at the scale of satellite sounding footprints. Dark gray pixels illustrate a single 8 pixel North–South swath through the middle of the gridcell. The black pixel represents a single satellite retrieval in the corner of the gridcell.

[13] The evaluated parameters ( $\Lambda$  and  $\mathbf{M}$ ) define the  $m \times m$  covariance matrix ( $\mathbf{V}_{\hat{y}}$ ) of the uncertainties of the  $X_{\text{CO}_2}$  signal at the resolution of the retrievals within each gridcell:

$$\mathbf{V}_{\hat{y}} = -\mathbf{X}\mathbf{M} + \mathbf{Q} - \mathbf{Q}\mathbf{S}\Lambda^T \quad (4)$$

The representation error, which is equal to the uncertainty associated with the estimated average (or block)  $X_{\text{CO}_2}$  within each gridcell, is evaluated by aggregating  $\mathbf{V}_{\hat{y}}$  as:

$$\sigma_{RE}^2 = [\mathbf{1}_m^T \mathbf{V}_{\hat{y}} \mathbf{1}_m] / m^2 \quad (5)$$

where  $\mathbf{1}_m$  is an  $m \times 1$  vector of ones. This estimated representation error, expressed as a variance, takes explicit account of both the spatial covariance structure of  $X_{\text{CO}_2}$  ( $\mathbf{Q}$ ), and the physical distribution (and redundancy) of retrievals within each gridcell.

## 2.2. X<sub>CO<sub>2</sub></sub> Spatial Variability

[14] To implement the method described in section 2.1, the spatial covariance of  $X_{\text{CO}_2}$  must be known. This covariance can be evaluated using aircraft measurements, or potentially satellite retrievals, in the geographic region of a gridcell. Alternately, as will be presented in this study, the covariance can be approximated based on model simulations of the global  $X_{\text{CO}_2}$  distribution. The spatial covariance information used in this study is based on work by *Alkhaled et al.* [2008], where the spatial variability of pressure-averaged dry-air mole fractions ( $X_{\text{CO}_2}$ ) was evaluated using simulations from the PCTM/GEOS-4 global chemistry and transport model run at a  $2^\circ$  latitude by  $2.5^\circ$  longitude resolution [*Kawa et al.*, 2004], as well as a second global model, a finer resolution regional model and aircraft measurements. *Alkhaled et al.* [2008] evaluated the spatial variability of  $X_{\text{CO}_2}$  as modeled by PCTM/GEOS-4 by

fitting the exponential covariance parameters  $\sigma_{reg}^2$  and  $L_{reg}$  (Section 2.1) in regions surrounding each gridcell. These parameters are used here to populate the covariance matrix ( $\mathbf{Q}$ ) as shown in equation (2).

## 2.3. Model Gridcell and Sampling Conditions

[15] In addition to  $X_{\text{CO}_2}$  variability over the sampled region, representation errors also depend on the satellite's retrieval footprint, the transport model resolution and the spatial distribution of retrievals within each gridcell.

[16] To demonstrate the proposed methodology, representation errors are quantified using a hypothetical transport model with  $1^\circ \times 1^\circ$  resolution and a  $3\text{km}^2$  retrieval footprint. Representation errors are evaluated assuming two spatial distributions of retrievals within each model gridcell, which represent idealized and adverse sampling conditions (Figure 1): gridcells are sampled assuming (1) a full North–South swath of retrievals in the middle of each gridcell, and (2) a single satellite retrieval at the corner of each gridcell. For illustration, the two sampling conditions are applied to all model gridcells, even at locations that would not be sampled due to the satellite track. The dimensions of the swath are representative of the sampling design of OCO, with 8 soundings across each swath. Each sounding is assumed to measure 2.4 km in latitude by 1.25 km in longitude.

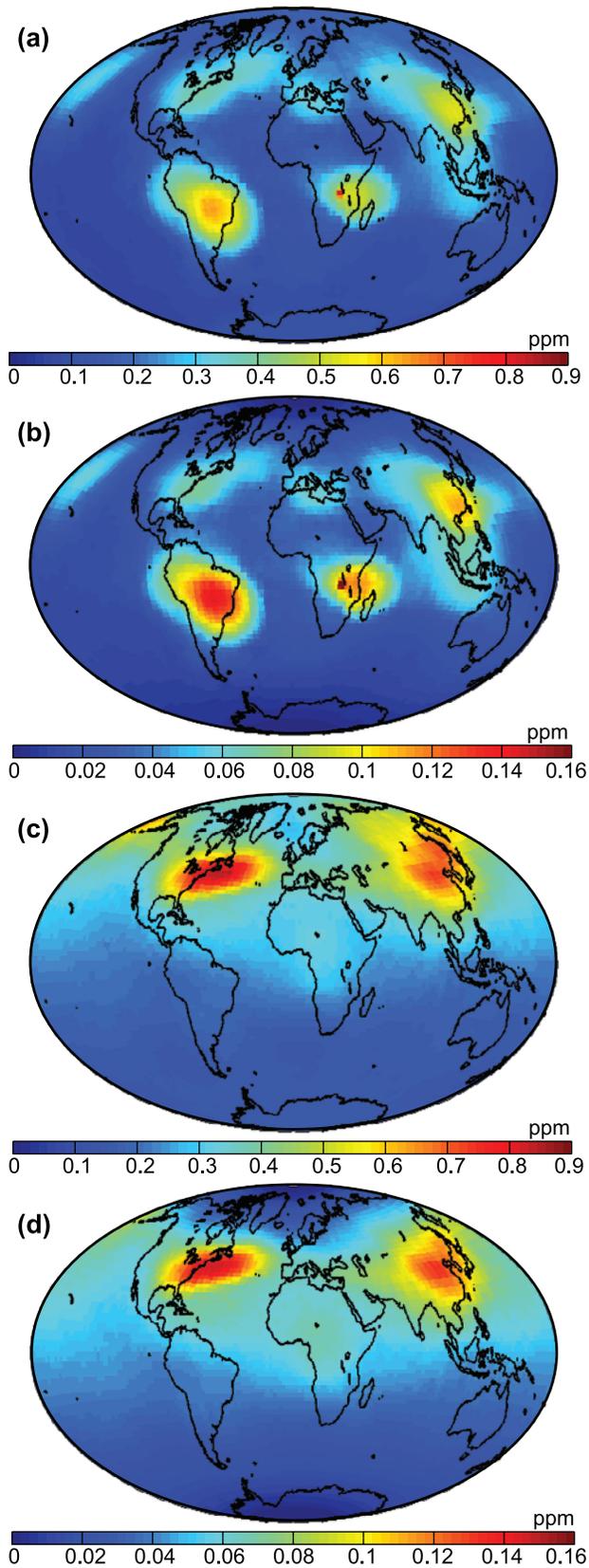
[17] To analyze the effects of the factors specifically controlling representation errors, no measurement error is included in the presented analysis ( $\mathbf{R} = \mathbf{0}$ ). For actual satellite retrievals, however, an accurate evaluation of the representation errors using the proposed method requires satellite measurement errors to be incorporated in equation 3.

[18] Although the example used here makes specific assumptions about model setup and satellite retrievals, the method can accommodate any transport model resolution, retrieval footprint, and retrieval distribution.

## 3. Results and Discussion

[19] To demonstrate the effects of seasonal changes in  $X_{\text{CO}_2}$  variability on representation errors, the presented method is applied for the months of January and July for both the swath and edge sampling described in Section 2.3. The regional spatial variability parameters evaluated by *Alkhaled et al.* [2008] range from  $0.24\text{ppm}^2$  to  $1.3\text{ppm}^2$  in January, and  $1.6\text{ppm}^2$  to  $9\text{ppm}^2$  in July. The shortest observed correlation lengths were 700 km in January, and 1800 km in July. The corresponding representation errors are presented in Figure 2 and show that: (1) representation errors are high over regions with high  $X_{\text{CO}_2}$  variability (see Figures 3, 4, and 5 of *Alkhaled et al.* [2008]), and (2) adverse sampling conditions increase representation errors even over areas with low  $X_{\text{CO}_2}$  variability.

[20] Seasonal changes in  $X_{\text{CO}_2}$  variability cause the location of maximum representation errors to vary seasonally. During the Northern Hemisphere (NH) summer, high representation errors occur over East Asia, Eastern North America and extend over the Atlantic Ocean, due to  $\text{CO}_2$  variability caused by North American fluxes (Figures 2c and 2d). During the NH winter, high representation errors occur over the Tropics and East Asia (Figures 2a and 2b).



**Figure 2.** Representation error assuming: one sounding per gridcell in (a) January and (c) July; one swath per gridcell in (b) January and (d) July.

Relatively high errors are also expected during the NH winter over Eastern North America. Representation errors are generally low over oceans and other continental areas.

[21] Figure 2 also demonstrates that the impact of the number of retrievals and their distribution within a gridcell is comparable to the impact of differences in  $X_{\text{CO}_2}$  variability. Figures 2a and 2c shows that for one satellite retrieval located in the corner of a gridcell, the representation errors range from 0.07 ppm to 0.82 ppm during the NH winter and 0.14 ppm to 0.86 ppm during summer. For a complete satellite swath in the middle of each gridcell, Figures 2b and 2d shows that the representation errors range up to 0.16 ppm during both the NH winter and summer.

[22] Results also show that representation errors are a function of gridcell area. The representation errors decrease for all model gridcells when moving from a single retrieval to a complete satellite swath, but this decrease is different for cells at the equator (large gridcell area) and for cells near the poles (small gridcell area).

[23] The presented method provides a flexible framework for accounting for the impact of geographic differences in  $X_{\text{CO}_2}$  variability and differences in the spatial distributions of retrieved soundings within gridcells. As such, the results presented here can be compared to representation errors reported in previous studies for cases involving similar sampling conditions. For example, *Miller et al.* [2007] estimated the representation error using  $X_{\text{CO}_2}$  simulated by a regional model over North America (NA), as described in Section 1. Representation errors were calculated for 1km and 16km grid resolutions for domains of 38km and 600km, respectively. Errors were found to be approximately 0.18 ppm for both the coarse and fine resolutions. Using a similar gridcell retrieval distribution and similar  $X_{\text{CO}_2}$  variability, the method presented in the current work produces similar results (0.12 ppm), as shown in Figure 2d over NA in July for a 10 km-wide swath and 3km<sup>2</sup> sounding footprint. A possible reason for the difference is the small height (7.2 km) of the  $X_{\text{CO}_2}$  column used by *Miller et al.* [2007] (i.e. higher  $X_{\text{CO}_2}$  variability) relative to the 48 km column used here. This comparison shows that for similar gridcell sampling conditions, region, and time, the presented method produces similar results, with the advantage that the actual gridcell mean need not be known to perform the analysis.

[24] Results can also be compared to those of *Corbin et al.* [2008] over NA and South America (SA) under swath sampling conditions. *Corbin et al.* [2008] evaluated the representation error for two model gridcell resolutions, 1 km and 5 km, and two grid sizes, 97 km for the fine resolution and 355 km to 450km for the coarse resolution. For these regions, August representation errors range from 0.09 ppm to 0.19 ppm over NA (results not shown), which are comparable to the values reported by *Corbin et al.* [2008] for the same month (0.06 ppm for the fine grid and 0.43 ppm for the coarse grid). Over SA, August representation errors range from 0.09 ppm to 0.20 ppm, which are also similar to *Corbin et al.* [2008] values of 0.21 ppm to 0.24 ppm for the fine and coarse grids, respectively. The advantage of the current method, however, is the ability to estimate representation errors without knowledge of all possible swath

means over a gridcell, which is required by *Corbin et al.* [2008] and will not be known for real satellite retrievals.

[25] *Gerbig et al.* [2003] and *Lin et al.* [2004] evaluated the spatial covariance of aircraft X<sub>CO<sub>2</sub></sub> measurements over NA and the Pacific Ocean, and used these covariances to produce statistical realizations of X<sub>CO<sub>2</sub></sub> at two model gridcell resolutions (5 km and 50 km) and a range of gridcell sizes (up to 1000 km). As discussed in Section 1, the representation error was then assumed to equal the average standard deviation of X<sub>CO<sub>2</sub></sub> values within all possible gridcells of the domain. In the case of a single retrieval per gridcell, the uncertainty associated with the inferred gridcell mean is equivalent to the variance of X<sub>CO<sub>2</sub></sub> at the retrieval resolution. Therefore, the representation errors reported by *Gerbig et al.* [2003] and *Lin et al.* [2004] are comparable to representation errors under adverse sampling conditions. Despite the mismatch between the sample and gridcell areas, the representation errors reported by *Gerbig et al.* [2003] and *Lin et al.* [2004] (0.5 ppm for NA and 0.25 ppm for the Pacific Ocean, as approximated from Figure 3 of *Lin et al.* [2004]) are comparable to the representation errors calculated here (Figures 2a and 2c), with the advantage that the approach presented here can accommodate any spatial distribution of samples within gridcells.

#### 4. Conclusions

[26] Representation errors occur due to the mismatch between the spatial footprint of X<sub>CO<sub>2</sub></sub> retrievals and the resolution of atmospheric transport models. The magnitude of these errors depends on the ability of retrieved soundings to capture the true X<sub>CO<sub>2</sub></sub> mean within model gridcells, which in turn depends on the number and spatial distribution of retrievals within model gridcells and the underlying spatial variability of X<sub>CO<sub>2</sub></sub> over the gridcell areas.

[27] This study introduces a geostatistical method for evaluating representation errors. Unlike previous studies, the method provides a statistical tool that quantifies grid-scale representation errors by linking the actual spatial distribution of retrievals within each gridcell and the regional X<sub>CO<sub>2</sub></sub> variability. The proposed method can evaluate errors associated with any model resolution and any satellite sounding footprint, as well as accounting for uncorrelated or correlated measurement errors. The X<sub>CO<sub>2</sub></sub> variability can be estimated using modeled X<sub>CO<sub>2</sub></sub> distributions, as was presented here, or could be inferred from actual satellite retrievals. The method does not require knowledge of the X<sub>CO<sub>2</sub></sub> distribution within gridcells at the resolution of the satellite footprint.

[28] The presented method was applied using spatial covariance information from *Alkhaled et al.* [2008], assuming a hypothetical model with 1° × 1° resolution and a sounding footprint representative of OCO soundings. Results show that representation errors vary spatially and temporally, as a function of seasonal and geographic changes in X<sub>CO<sub>2</sub></sub> variability, and the spatial distribution of satellite retrievals within each gridcell.

[29] Although this study focused on spatial representation errors, temporal X<sub>CO<sub>2</sub></sub> variability would also contribute to

representation errors if retrievals taken across multiple days were used jointly to estimate X<sub>CO<sub>2</sub></sub> for a given location and time. Extending the presented method to include temporal variability will be the topic of future work.

[30] **Acknowledgments.** This research was partially performed for the Orbiting Carbon Observatory Project at Caltech-JPL, under a contract with NASA. Additional support was provided through NASA grant NNX08AJ92G, and the Kuwait University Scholarship Committee. The PCTM work was enabled by G. J. Collatz and Z. Zhu, and supported by NASA Carbon Cycle Science.

#### References

- Alkhaled, A., A. M. Michalak, S. R. Kawa, S. Olsen, and J. W. Wang (2008), A global evaluation of the regional spatial variability of column integrated CO<sub>2</sub> Distributions, *J. Geophys. Res.*, doi:10.1029/2007JD009693, in press.
- Chevallier, F., F.-M. Breon, and P. J. Rayner (2007), Contribution of the Orbiting Carbon Observatory to the estimation of CO<sub>2</sub> sources and sinks: Theoretical study in a variational data assimilation framework, *J. Geophys. Res.*, 112, D09307, doi:10.1029/2006JD007375.
- Chilès, J. P., and P. Delfiner (1999), *Geostatistics: Modeling Spatial Uncertainty*, John Wiley, Hoboken, N. J.
- Corbin, K., A. S. Denning, J.-W. Wang, L. Lu, and I. T. Baker (2008), Possible representation errors in inversions of satellite CO<sub>2</sub> retrievals, *J. Geophys. Res.*, 113, D02301, doi:10.1029/2007JD008716.
- Engelen, R. J., A. S. Denning, K. R. Gurney, and TransCom3 modelers (2002), On error estimation in atmospheric CO<sub>2</sub> inversions, *J. Geophys. Res.*, 107(D22), 4635, doi:10.1029/2002JD002195.
- Enting, I. G. (2002), *Inverse Problems in Atmospheric Constituent Transport*, Cambridge Univ. Press, New York.
- Gerbig, C., J. C. Lin, S. C. Wofsy, B. C. Daube, A. E. Andrews, B. B. Stephens, P. S. Bakwin, and C. A. Grainger (2003), Toward constraining regional-scale fluxes of CO<sub>2</sub> with atmospheric observations over a continent: 1. Observed spatial variability from airborne platforms, *J. Geophys. Res.*, 108(D24), 4756, doi:10.1029/2002JD003018.
- Houweling, S., F. M. Breon, I. Aben, C. Rodenbeck, M. Gloor, M. Heimann, and P. Ciais (2004), Inverse modeling of CO<sub>2</sub> sources and sinks using satellite data: A synthetic inter-comparison of measurement techniques and their performance as a function of space and time, *Atmos. Chem. Phys.*, 4, 523–538.
- Kawa, S. R., D. J. Erickson, S. Pawson, and Z. Zhu (2004), Global CO<sub>2</sub> transport simulations using meteorological data from the NASA data assimilation system, *J. Geophys. Res.*, 109, D18312, doi:10.1029/2004JD004554.
- Lin, J. C., C. Gerbig, B. C. Daube, S. C. Wofsy, A. E. Andrews, S. A. Vay, and B. E. Anderson (2004), An empirical analysis of the spatial variability of atmospheric CO<sub>2</sub>: Implications for inverse analyses and space-borne sensors, *Geophys. Res. Lett.*, 31, L23104, doi:10.1029/2004GL020957.
- Michalak, A. M., A. Hirsch, L. Bruhwiler, K. R. Gurney, W. Peters, and P. P. Tans (2005), Maximum likelihood estimation of covariance parameters for Bayesian atmospheric trace gas surface flux inversions, *J. Geophys. Res.*, 110, D24107, doi:10.1029/2005JD005970.
- Miller, C. E., et al. (2007), Precision requirements for space-based X<sub>CO<sub>2</sub></sub> data, *J. Geophys. Res.*, 112, D10314, doi:10.1029/2006JD007659.
- Rayner, P. J., and D. M. O'Brien (2001), The utility of remotely sensed CO<sub>2</sub> concentration data in surface source inversions, *Geophys. Res. Lett.*, 28, 175–178.
- Rödenbeck, C., S. Houweling, M. Gloor, and M. Heimann (2003), CO<sub>2</sub> flux history 1982–2001 inferred from atmospheric data using a global inversion of atmospheric transport, *Atmos. Chem. Phys.*, 3, 1919–1964.
- van der Molen, M. K., and A. J. Dolman (2007), Regional carbon fluxes and the effect of topography on the variability of atmospheric CO<sub>2</sub>, *J. Geophys. Res.*, 112, D01104, doi:10.1029/2006JD007649.

A. A. Alkhaled and A. M. Michalak, Department of Civil and Environmental Engineering, University of Michigan, 183 EWRE Building, Ann Arbor, MI 48109, USA. (alanood@umich.edu; amichala@umich.edu)  
S. R. Kawa, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA.